

The Three-Dimensional Gauss Algorithm Is Strongly Convergent Almost Everywhere

D. M. Hardcastle

CONTENTS

- 1. Introduction
 - 2. The Three-Dimensional Gauss Transformation
 - 3. Strategy of the Proof
 - 4. Numerical Estimation
 - 5. Numerical Results
 - 6. Discussion
- References

A proof that the three-dimensional Gauss algorithm is strongly convergent almost everywhere is given. This algorithm is equivalent to Brun's algorithm and to the modified Jacobi-Perron algorithm considered by Podsypanin and Schweiger. The proof involves the rigorous computer assisted estimation of the largest Lyapunov exponent of a cocycle associated to the algorithm. To the best of my knowledge, this is the first proof of almost everywhere strong convergence of a Jacobi-Perron type algorithm in dimension greater than two.

1. INTRODUCTION

Multidimensional continued fraction algorithms have been widely studied since the nineteenth century. They were first introduced by Jacobi who considered a generalisation of the famous one-dimensional continued fraction algorithm to two dimensions [Jacobi 1868]. Jacobi's algorithm is now known as the Jacobi-Perron algorithm (JPA), in recognition of the fundamental work of Perron who defined the algorithm in arbitrary dimension and proved its convergence [Perron 1907]. Other famous algorithms include Brun's [Brun 57], Selmer's [Selmer 61] and the modified Jacobi-Perron algorithm considered by Podsypanin [Podsypanin 77] and Schweiger [Schweiger 79].

For a given irrational vector, $(\omega_1, \dots, \omega_d) \in [0, 1]^d \setminus \mathbb{Q}^d$, all multidimensional continued fraction algorithms produce a sequence of rational vectors $(p_1(n)/q(n), \dots, p_d(n)/q(n))$ which converges to $(\omega_1, \dots, \omega_d)$. The algorithms mentioned above are known to be convergent in the weak sense, i.e.

$$\lim_{n \rightarrow \infty} \left\| (\omega_1, \dots, \omega_d) - \left(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)} \right) \right\| = 0.$$

The approximations are said to be strongly convergent if

$$\lim_{n \rightarrow \infty} \|q(n)(\omega_1, \dots, \omega_d) - (p_1(n), \dots, p_d(n))\| = 0. \quad (1-1)$$

2000 AMS Subject Classification: Primary 11J70; Secondary 11K50

Keywords: Multidimensional continued fractions, Brun's algorithm, Jacobi-Perron algorithm, strong convergence, Lyapunov exponents

It is believed that the JPA, the modified JPA and the algorithms of Brun and Selmer are strongly convergent almost everywhere, i.e. for Lebesgue almost all $(\omega_1, \dots, \omega_d) \in [0, 1]^d \setminus \mathbb{Q}^d$ equation (1–1) holds. However, at the present time, the only rigorous proofs of strong convergence are for two-dimensional algorithms. In fact, strong convergence of the two-dimensional JPA is a consequence of the results of Paley and Ursell [Paley, Ursell 30]. This observation was first made by K. Khanin [Khanin 92] (see also [Schweiger 96]). The modified JPA and Brun’s algorithm have also been proven to be strongly convergent almost everywhere in two dimensions (see [Meester 98] and [Schratzberger 98]).

In the early 1990s the significance of the Lyapunov exponents for the convergence properties of multidimensional continued fraction algorithms was noted, independently, by D. Kosygin [Kosygin 91] and P. Baldwin [Baldwin 92a, Baldwin 92b]. A d -dimensional continued fraction algorithm has $d + 1$ Lyapunov exponents $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{d+1}$. The approximations produced by an algorithm are exponentially strongly convergent almost everywhere if and only if $\lambda_2 < 0$. In [Ito et al. 93], Ito, Keane and Ohtsuki introduced a cocycle with largest Lyapunov exponent λ_2 . They used this cocycle to give a computer assisted proof of almost everywhere exponential strong convergence of the two-dimensional modified JPA.

In [Hardcastle, Khanin 00], K. Khanin and I introduced a method which, in principle, can be used to prove almost everywhere strong convergence in arbitrary dimension. This scheme requires a good knowledge of the invariant measure of the endomorphism associated to the algorithm. The method was applied to the ordered JPA which is equivalent to the modified JPA and Brun’s algorithm. The ordered JPA is particularly suitable for study by this scheme because the invariant density is known explicitly.

In [Hardcastle, Khanin 02] the method of Ito et al. was discussed in arbitrary dimension. In particular it was explained how the problem of proving almost everywhere exponential strong convergence of the ordered JPA can be reduced to a finite number of calculations. In this paper, these calculations are carried out for the three-dimensional ordered JPA. This results in a computer assisted proof of almost everywhere strong convergence of the three-dimensional ordered JPA.

Note that I use the term “proof” here, despite the fact that I do not attempt to control round-off errors. I will leave the issue of whether the term “proof” is appropriate to the individual reader.

Both in [Hardcastle, Khanin 02] and in this paper the ordered JPA is called the d -dimensional Gauss algorithm. This name was introduced in [Hardcastle, Khanin 01] where it was suggested that the ordered JPA should be considered to be the most natural generalisation of the one-dimensional Gauss transformation.

2. THE THREE-DIMENSIONAL GAUSS TRANSFORMATION

This section begins with a description of the three-dimensional Gauss algorithm. The ergodic properties of the algorithm are then discussed and we explain the significance of the Lyapunov exponents for the convergence properties of the algorithm.

Let $\Delta^3 = \{(\omega_1, \omega_2, \omega_3) \in [0, 1]^3 : \omega_1 \geq \omega_2 \geq \omega_3\}$. Define $T : \Delta^3 \rightarrow \Delta^3$ by

$$T(\omega_1, \omega_2, \omega_3) = \begin{cases} \left(\left\{ \frac{1}{\omega_1} \right\}, \frac{\omega_2}{\omega_1}, \frac{\omega_3}{\omega_1} \right) & \text{if } \left\{ \frac{1}{\omega_1} \right\} > \frac{\omega_2}{\omega_1}; \\ \left(\frac{\omega_2}{\omega_1}, \left\{ \frac{1}{\omega_1} \right\}, \frac{\omega_3}{\omega_1} \right) & \text{if } \frac{\omega_2}{\omega_1} > \left\{ \frac{1}{\omega_1} \right\} > \frac{\omega_3}{\omega_1}; \\ \left(\frac{\omega_2}{\omega_1}, \frac{\omega_3}{\omega_1}, \left\{ \frac{1}{\omega_1} \right\} \right) & \text{if } \frac{\omega_3}{\omega_1} > \left\{ \frac{1}{\omega_1} \right\}. \end{cases} \quad (2-1)$$

In this formula, $\{x\}$ denotes the fractional part of a real number x , i.e. $\{x\} = x - [x]$ where $[x]$ denotes the integer part of x .

Definition 2.1. The transformation $T : \Delta^3 \rightarrow \Delta^3$ is called the *three-dimensional Gauss transformation*.

Two numbers are naturally associated to the process of calculating $T(\omega) = T(\omega_1, \omega_2, \omega_3)$: the integer part of $1/\omega_1$ which we denote by $m(\omega)$, i.e. $m(\omega) = [1/\omega_1]$, and the position in which $\{1/\omega_1\} = 1/\omega_1 - m(\omega)$ is placed, which we denote by $j(\omega)$, i.e. the $j(\omega)$ th coordinate of $T(\omega)$ is $\{1/\omega_1\}$.

Define $\Delta_{(m,j)} = \{\omega \in \Delta^3 : m(\omega) = m, j(\omega) = j\}$. It is easy to check that $T|_{\Delta_{(m,j)}}$ is injective and the image of $\Delta_{(m,j)}$ under T is the whole of Δ^3 . The inverse of $T|_{\Delta_{(m,j)}}$, which we denote $T_{(m,j)}^{-1}$, is given by

$$T_{(m,j)}^{-1}(\omega_1, \omega_2, \omega_3) = \begin{cases} \left(\frac{1}{m + \omega_1}, \frac{\omega_2}{m + \omega_1}, \frac{\omega_3}{m + \omega_1} \right) & \text{if } j = 1; \\ \left(\frac{1}{m + \omega_2}, \frac{\omega_1}{m + \omega_2}, \frac{\omega_3}{m + \omega_2} \right) & \text{if } j = 2; \\ \left(\frac{1}{m + \omega_3}, \frac{\omega_1}{m + \omega_3}, \frac{\omega_2}{m + \omega_3} \right) & \text{if } j = 3. \end{cases}$$

For each $(m, j) \in \mathbb{N} \times \{1, 2, 3\}$ we define a matrix $\tilde{A}_{(m,j)} \in GL(4, \mathbb{Z})$:

$$\tilde{A}_{(m,1)} = \begin{pmatrix} m & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \tilde{A}_{(m,2)} = \begin{pmatrix} m & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$\tilde{A}_{(m,3)} = \begin{pmatrix} m & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The matrices $\tilde{A}_{(m,j)}$ give the action of $T_{(m,j)}^{-1}$ on rational vectors:

$$T_{(m,j)}^{-1} \left(\frac{p_1}{q}, \frac{p_2}{q}, \frac{p_3}{q} \right) = \left(\frac{\tilde{p}_1}{\tilde{q}}, \frac{\tilde{p}_2}{\tilde{q}}, \frac{\tilde{p}_3}{\tilde{q}} \right)$$

if and only if $\begin{pmatrix} \tilde{q} \\ \tilde{p}_1 \\ \tilde{p}_2 \\ \tilde{p}_3 \end{pmatrix} = \tilde{A}_{(m,j)} \begin{pmatrix} q \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}.$

To form simultaneous rational approximations to a vector $\omega \in \Delta^3$ we consider the orbit of ω under T :

$$\omega \xrightarrow{T} T\omega \xrightarrow{T} \dots \xrightarrow{T} T^{n-1}\omega.$$

To this trajectory we associate the sequence $(m_1, j_1), \dots, (m_n, j_n)$ where

$$m_i = m(T^{i-1}\omega), \quad j_i = j(T^{i-1}\omega).$$

Define

$$\tilde{C}_n(\omega) = \tilde{A}_{(m_1, j_1)} \tilde{A}_{(m_2, j_2)} \cdots \tilde{A}_{(m_n, j_n)}. \quad (2-2)$$

The matrix $\tilde{C}_n(\omega)$ gives the n^{th} approximation to ω . More precisely, let

$$\tilde{C}_n(\omega) = \begin{pmatrix} q(n, 0) & q(n, 1) & q(n, 2) & q(n, 3) \\ p_1(n, 0) & p_1(n, 1) & p_1(n, 2) & p_1(n, 3) \\ p_2(n, 0) & p_2(n, 1) & p_2(n, 2) & p_2(n, 3) \\ p_3(n, 0) & p_3(n, 1) & p_3(n, 2) & p_3(n, 3) \end{pmatrix}.$$

Denote

$$\frac{\mathbf{p}(n, i)}{q(n, i)} = \left(\frac{p_1(n, i)}{q(n, i)}, \frac{p_2(n, i)}{q(n, i)}, \frac{p_3(n, i)}{q(n, i)} \right), \quad 0 \leq i \leq 3. \quad (2-3)$$

We will consider $\mathbf{p}(n, 0)/q(n, 0)$ as the n^{th} approximation to ω . Denote

$$\tau_n = \frac{\mathbf{p}(n, 0)}{q(n, 0)} = \left(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)} \right). \quad (2-4)$$

Definition 2.2. A sequence of rational vectors $\tau_n = (p_1(n)/q(n), \dots, p_d(n)/q(n))$ is *exponentially strongly convergent to ω* if there exist constants $k > 0$, $\alpha > 0$ such that

$$\|q(n)\omega - (p_1(n), \dots, p_d(n))\| \leq kq(n)^{-\alpha}.$$

The purpose of this paper is to prove that the three-dimensional Gauss algorithm is exponentially strongly convergent almost everywhere, i.e. for almost all $\omega \in \Delta^3$, the sequence τ_n defined by (2-4) is exponentially strongly convergent to ω .

It will be convenient to present (2-2) in a slightly different fashion. Let $A_{(m,j)} = (\tilde{A}_{(m,j)})^t$ where A^t denotes the transpose of a matrix A . Define a matrix-valued function on Δ^3 by

$$A(\omega) = A_{(m(\omega), j(\omega))}.$$

Denote

$$C_n(\omega) = A(T^{n-1}\omega) \cdots A(T\omega)A(\omega).$$

Then

$$C_n(\omega) = (\tilde{C}_n(\omega))^t = A_{(m_n, j_n)} \cdots A_{(m_2, j_2)} A_{(m_1, j_1)}.$$

We now discuss the ergodic properties of the map T . Define

$$\begin{aligned} \rho(\omega) &= \sum_{\pi \in S_3} \frac{1}{1 + \omega_{\pi(1)}} \frac{1}{1 + \omega_{\pi(1)} + \omega_{\pi(2)}} \\ &\quad \times \frac{1}{1 + \omega_{\pi(1)} + \omega_{\pi(2)} + \omega_{\pi(3)}} \\ &= \frac{1}{1 + \omega_1 + \omega_2 + \omega_3} \left(\frac{1}{(1 + \omega_1)(1 + \omega_1 + \omega_2)} \right. \\ &\quad + \frac{1}{(1 + \omega_1)(1 + \omega_1 + \omega_3)} + \frac{1}{(1 + \omega_2)(1 + \omega_2 + \omega_1)} \\ &\quad + \frac{1}{(1 + \omega_2)(1 + \omega_2 + \omega_3)} + \frac{1}{(1 + \omega_3)(1 + \omega_3 + \omega_1)} \\ &\quad \left. + \frac{1}{(1 + \omega_3)(1 + \omega_3 + \omega_2)} \right) \end{aligned} \quad (2-5)$$

where S_3 is the group of permutations of $\{1, 2, 3\}$. Let $K = \int_{\Delta^3} \rho(\omega) d\omega$. It is easy to check that the probability measure μ defined by

$$\mu(X) = \frac{1}{K} \int_X \rho(\omega) d\omega, \quad X \text{ a Borel subset of } \Delta^d,$$

is T -invariant and ergodic (see [Schweiger 79]). This measure is the unique absolutely continuous T -invariant probability measure.

The endomorphism T , the matrix-valued function A and the invariant measure μ together form a cocycle which we denote (T, A, μ) . This cocycle is integrable, i.e.

$$\int_{\Delta^3} \log(\max(\|A(\omega)\|, 1)) \mu(d\omega) < \infty.$$

Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ be the Lyapunov exponents of the cocycle (T, A, μ) (see [Oseledets 68]).

The following result relates the convergence properties of the algorithm to its Lyapunov exponents.

Theorem 2.3.

(i) *The largest Lyapunov exponent λ_1 is strictly positive and simple.*

(ii) *For almost all $\omega \in \Delta^3$*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log q(n) = \lambda_1.$$

(iii) *The sequence τ_n is exponentially strongly convergent to ω for almost all ω if and only if $\lambda_2 < 0$.*

This theorem, which is based on the work of Lagarias [Lagarias 93], was proved in [Hardcastle, Khanin 00].

3. STRATEGY OF THE PROOF

By Theorem 2.3, in order to prove almost everywhere strong convergence, it is enough to show that $\lambda_2 < 0$. It is difficult to estimate the exponent λ_2 directly, so we consider a new cocycle (T, D, μ) whose largest Lyapunov exponent $\lambda_1(D)$ is λ_2 . Our strategy will be to estimate $\lambda_1(D)$ by use of the Subadditive Ergodic Theorem.

Define

$$D_1 = \begin{pmatrix} -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; D_2 = \begin{pmatrix} 0 & 1 & 0 \\ -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 0 & 1 \end{pmatrix};$$

and

$$D_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\omega_1 & -\omega_2 & -\omega_3 \end{pmatrix}.$$

Let $D(\omega) = D_{j(\omega)}$. Also define

$$D_n(\omega) = D(T^{n-1}\omega) \cdots D(T\omega)D(\omega).$$

Let $\lambda_1(D)$ denote the largest Lyapunov exponent of the cocycle (T, D, μ) .

Lemma 3.1. $\lambda_1(D) = \lambda_2$.

In [Hardcastle, Khanin 02] we gave a proof of this lemma and a description of the construction of the matrices $D(\omega)$. Our construction was based on an observation made by Lagarias [Lagarias 93], but matrices similar to $D(\omega)$ have also been considered by Ito, Keane and Ohtsuki [Ito et al. 93, Ito et al. 96].

The following lemma is an immediate consequence of the Subadditive Ergodic Theorem [Kingman 68].

Lemma 3.2. $\lambda_1(D) = \inf_{n \in \mathbb{N}} \frac{1}{n} \int_{\Delta^3} \log \|D_n(\omega)\| \mu(d\omega)$.

Combining Lemmas 3.1 and 3.2, and Theorem 2.3 we have:

Theorem 3.3. *The three-dimensional Gauss algorithm is exponentially strongly convergent almost everywhere if and only if there exists $n \in \mathbb{N}$ such that*

$$\frac{1}{n} \int_{\Delta^3} \log \|D_n(\omega)\| \mu(d\omega) < 0. \tag{3-1}$$

We will give a computer assisted proof that condition (3-1) holds for $n = 8$. The following results, which were proved in [Hardcastle, Khanin 02], will be useful when we estimate the integral in (3-1).

We first show that the matrix $D_n(\omega)$ can be expressed in terms of ω and its approximations.

Proposition 3.4. *For all $n \in \mathbb{N}$*

$$D_n(\omega) = \begin{pmatrix} q(n, 1)\omega_1 - p_1(n, 1) & q(n, 1)\omega_2 - p_2(n, 1) & q(n, 1)\omega_3 - p_3(n, 1) \\ q(n, 2)\omega_1 - p_1(n, 2) & q(n, 2)\omega_2 - p_2(n, 2) & q(n, 2)\omega_3 - p_3(n, 2) \\ q(n, 3)\omega_1 - p_1(n, 3) & q(n, 3)\omega_2 - p_2(n, 3) & q(n, 3)\omega_3 - p_3(n, 3) \end{pmatrix}.$$

Recall that in Section 2 we defined

$$\Delta_{(m,j)} = \{\omega \in \Delta^3 : m(\omega) = m, j(\omega) = j\}.$$

These sets form a Markov partition for T . Let $\Delta_{(m_1, j_1), \dots, (m_n, j_n)}$ denote an element of the n^{th} level of the Markov partition:

$$\Delta_{(m_1, j_1), \dots, (m_n, j_n)} = \{\omega \in \Delta^3 : m(T^{i-1}\omega) = m_i, j(T^{i-1}\omega) = j_i \text{ for } 1 \leq i \leq n\}.$$

The following proposition describes how the vertices of the simplex $\Delta_{(m_1, j_1), \dots, (m_n, j_n)}$ can be calculated.

Proposition 3.5. *Let*

$$\begin{aligned} V_{(m_1, j_1), \dots, (m_n, j_n)} &= (v_{il})_{1 \leq i, l \leq 4} \\ &= \tilde{A}_{(m_1, j_1)} \tilde{A}_{(m_2, j_2)} \cdots \tilde{A}_{(m_n, j_n)} V_0 \end{aligned}$$

where

$$V_0 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then the vertices of $\Delta_{(m_1, j_1), \dots, (m_n, j_n)}$ are

$$\mathbf{v}_i = \left(\frac{v_{2i}}{v_{1i}}, \frac{v_{3i}}{v_{1i}}, \frac{v_{4i}}{v_{1i}} \right), \quad 1 \leq i \leq 4.$$

The next result explains how the maximum value of $\|D_n(\boldsymbol{\omega})\|$ over a simplex can be calculated. Let $\|\cdot\|$ denote an arbitrary norm on \mathbb{R}^3 . Then the corresponding norm of a linear operator $D : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is given by

$$\|D\| = \sup_{\mathbf{v} \in \mathbb{R}^3 \setminus \{0\}} \frac{\|D\mathbf{v}\|}{\|\mathbf{v}\|}.$$

Proposition 3.6. *The maximum value of $\log \|D_n(\boldsymbol{\omega})\|$ over a simplex $\Delta \subseteq \Delta_{(m_1, j_1), \dots, (m_n, j_n)}$ occurs at a vertex of Δ , i.e.*

$$\max_{\boldsymbol{\omega} \in \Delta} \log \|D_n(\boldsymbol{\omega})\| = \max_{1 \leq i \leq 4} \log \|D_n(\mathbf{v}_i)\|$$

where $\mathbf{v}_1, \dots, \mathbf{v}_4$ are the vertices of Δ .

Proof: See Corollary 4.7 of [Hardcastle, Khanin 02]. \square

We now consider how we can calculate an upper and a lower bound for the density ρ over a simplex.

Proposition 3.7. *The maximum value of ρ over a simplex $\Delta \subseteq \Delta^3$ occurs at one of the vertices of Δ , i.e.*

$$\max_{\boldsymbol{\omega} \in \Delta} \rho(\boldsymbol{\omega}) = \rho(\mathbf{v})$$

where \mathbf{v} is a vertex of Δ .

Proof: See Corollary 4.11 of [Hardcastle, Khanin 02]. \square

Corollary 3.8. *If $\Delta \subseteq \Delta^3$ is a simplex then*

$$\mu(\Delta) \leq \frac{1}{K} \left(\max_{1 \leq i \leq 4} \rho(\mathbf{v}_i) \right) \text{vol}(\Delta)$$

where $\mathbf{v}_1, \dots, \mathbf{v}_4$ are the vertices of Δ and vol denotes three-dimensional Lebesgue measure.

Define functions $s_1, s_2, s_3, s_{12}, s_{13}, s_{23}, s_{123} : \Delta^3 \rightarrow \mathbb{R}$ by

$$\begin{aligned} s_1(\boldsymbol{\omega}) &= \omega_1, & s_2(\boldsymbol{\omega}) &= \omega_2, & s_3(\boldsymbol{\omega}) &= \omega_3, \\ s_{12}(\boldsymbol{\omega}) &= \omega_1 + \omega_2, & s_{13}(\boldsymbol{\omega}) &= \omega_1 + \omega_3, \\ s_{23}(\boldsymbol{\omega}) &= \omega_2 + \omega_3, & s_{123}(\boldsymbol{\omega}) &= \omega_1 + \omega_2 + \omega_3. \end{aligned}$$

Then

$$\begin{aligned} \rho(\boldsymbol{\omega}) &= \frac{1}{1 + s_{123}(\boldsymbol{\omega})} \\ &\times \left(\frac{1}{1 + s_1(\boldsymbol{\omega})} \left(\frac{1}{1 + s_{12}(\boldsymbol{\omega})} + \frac{1}{1 + s_{13}(\boldsymbol{\omega})} \right) \right. \\ &\quad + \frac{1}{1 + s_2(\boldsymbol{\omega})} \left(\frac{1}{1 + s_{12}(\boldsymbol{\omega})} + \frac{1}{1 + s_{23}(\boldsymbol{\omega})} \right) \\ &\quad \left. + \frac{1}{1 + s_3(\boldsymbol{\omega})} \left(\frac{1}{1 + s_{13}(\boldsymbol{\omega})} + \frac{1}{1 + s_{23}(\boldsymbol{\omega})} \right) \right). \end{aligned}$$

Lemma 3.9. *Let $\Delta \subseteq \Delta^3$ be a simplex with vertices $\mathbf{v}_1, \dots, \mathbf{v}_4$. The maximum value of each function s_1, s_2, \dots, s_{123} is attained at a vertex of Δ , i.e.*

$$\max_{\boldsymbol{\omega} \in \Delta} s_l(\boldsymbol{\omega}) = \max_{1 \leq i \leq 4} s_l(\mathbf{v}_i)$$

where $l = 1, 2, 3, 12, 13, 23, 123$.

Let $\bar{s}_1 = \max_{1 \leq i \leq 4} s_1(\mathbf{v}_i)$ and similarly for $\bar{s}_2, \bar{s}_3, \dots, \bar{s}_{123}$.

Corollary 3.10. *For $\boldsymbol{\omega} \in \Delta$*

$$\begin{aligned} \rho(\boldsymbol{\omega}) &\geq \frac{1}{1 + \bar{s}_{123}} \left(\frac{1}{1 + \bar{s}_1} \left(\frac{1}{1 + \bar{s}_{12}} + \frac{1}{1 + \bar{s}_{13}} \right) \right. \\ &\quad + \frac{1}{1 + \bar{s}_2} \left(\frac{1}{1 + \bar{s}_{12}} + \frac{1}{1 + \bar{s}_{23}} \right) \\ &\quad \left. + \frac{1}{1 + \bar{s}_3} \left(\frac{1}{1 + \bar{s}_{13}} + \frac{1}{1 + \bar{s}_{23}} \right) \right). \end{aligned}$$

Let the above lower bound for $\rho(\boldsymbol{\omega})$ be denoted $\underline{\rho}(\Delta)$. Then we have:

Corollary 3.11. $\mu(\Delta) \geq \frac{1}{K} \underline{\rho}(\Delta) \text{vol}(\Delta)$.

The following two lemmas will allow us to obtain a global bound for $\|D_n(\boldsymbol{\omega})\|$.

Lemma 3.12. *For any $\boldsymbol{\omega} \in \Delta^3$, $\|D(T\boldsymbol{\omega})D(\boldsymbol{\omega})\| \leq \sqrt{7}$.*

Proof: Suppose $j(\boldsymbol{\omega}) = 1$, so that

$$T(\boldsymbol{\omega}) = T(\omega_1, \omega_2, \omega_3) = \left(\frac{1}{\omega_1} - m_1, \frac{\omega_2}{\omega_1}, \frac{\omega_3}{\omega_1} \right).$$

If $j(T\omega) = 1$ then

$$\begin{aligned} D(T\omega)D(\omega) &= \begin{pmatrix} m_1 - \frac{1}{\omega_1} & -\frac{\omega_2}{\omega_1} & -\frac{\omega_3}{\omega_1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 - m_1\omega_1 & -m_1\omega_2 & -m_1\omega_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{3-2}$$

Suppose $\|\mathbf{x}\| = \|(x_1, x_2, x_3)\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = 1$. Then $\|D(T\omega)D(\omega)\mathbf{x}\|^2$

$$\begin{aligned} &= \left\| \begin{pmatrix} 1 - m_1\omega_1 & -m_1\omega_2 & -m_1\omega_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right\|^2 \\ &= ((1 - m_1\omega_1)x_1 - m_1\omega_2x_2 - m_1\omega_3x_3)^2 + x_2^2 + x_3^2 \\ &\leq ((1 - m_1\omega_1)^2 + m_1^2\omega_2^2 + m_1^2\omega_3^2)(x_1^2 + x_2^2 + x_3^2) \\ &\quad + 1 - x_1^2 \\ &\leq 3 + 1 = 4, \end{aligned}$$

since $0 \leq \omega_3 \leq \omega_2 \leq \omega_1 \leq \frac{1}{m_1}$. Thus $\|D(T\omega)D(\omega)\| \leq 2$. Now if $j(T\omega) = 2$ or 3 then $D(T\omega)D(\omega)$ will be the same as (3-2) up to a change in the order of the rows. Thus, for any ω such that $j(\omega) = 1$, $\|D(T\omega)D(\omega)\| \leq 2$.

Now suppose that $j(\omega) = 2$ so that

$$T(\omega) = T(\omega_1, \omega_2, \omega_3) = \begin{pmatrix} \omega_2 & 1 \\ \omega_1 & -m_1, \frac{\omega_3}{\omega_1} \end{pmatrix}.$$

Then if $j(T\omega) = 1$

$$\begin{aligned} D(T\omega)D(\omega) &= \begin{pmatrix} -\frac{\omega_2}{\omega_1} & m_1 - \frac{1}{\omega_1} & -\frac{\omega_3}{\omega_1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 - m_1\omega_1 & -m_1\omega_2 & -m_1\omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

So, if $x_1^2 + x_2^2 + x_3^2 = 1$ then

$$\begin{aligned} &\|D(T\omega)D(\omega)\mathbf{x}\|^2 \\ &= \left\| \begin{pmatrix} 1 - m_1\omega_1 & -m_1\omega_2 & -m_1\omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right\|^2 \\ &= ((1 - m_1\omega_1)x_1 - m_1\omega_2x_2 - m_1\omega_3x_3)^2 \\ &\quad + (\omega_1x_1 + \omega_2x_2 + \omega_3x_3) + x_3^2 \\ &\leq ((1 - m_1\omega_1)^2 + m_1^2\omega_2^2 + m_1^2\omega_3^2)(x_1^2 + x_2^2 + x_3^2) \\ &\quad + (\omega_1^2 + \omega_2^2 + \omega_3^2) + 1 - x_1^2 - x_2^2 \\ &\leq 3 + 3 + 1 = 7. \end{aligned}$$

Thus

$$\|D(T\omega)D(\omega)\| \leq \sqrt{7}.$$

If $j(T\omega) = 2$ or 3 then the same estimate holds, and it can easily be checked that if $j(\omega) = 3$ then the same estimate is valid. \square

Lemma 3.13. For any $\omega \in \Delta^3$

$$\begin{aligned} \|D(\omega)\|^2 &= \frac{1}{2} \left(1 + \omega_1^2 + \omega_2^2 + \omega_3^2 + \sqrt{(\omega_1^2 + \omega_2^2 + \omega_3^2 + 1)^2 - 4\omega_1^2} \right). \end{aligned}$$

Proof: This is an easy calculation. \square

Corollary 3.14. If $m(\omega) \geq m$ then

$$\|D(\omega)\|^2 \leq \frac{1}{2} \left(1 + \frac{3}{m^2} + \sqrt{\frac{9}{m^4} + \frac{4}{m^2} + 1} \right).$$

Proof: By the lemma,

$$\begin{aligned} \|D(\omega)\|^2 &= \frac{1}{2} \left(1 + \omega_1^2 + \omega_2^2 + \omega_3^2 \right. \\ &\quad \left. + \sqrt{\omega_1^4 + \omega_2^4 + \omega_3^4 + 2(\omega_1^2\omega_2^2 + \omega_1^2\omega_3^2 + \omega_2^2\omega_3^2 - \omega_1^2 + \omega_2^2 + \omega_3^2) + 1} \right). \end{aligned}$$

If $m(\omega) \geq m$ then $\omega_3 \leq \omega_2 \leq \omega_1 \leq \frac{1}{m}$, so

$$\begin{aligned} \|D(\omega)\|^2 &\leq \frac{1}{2} \left(1 + \frac{3}{m^2} + \sqrt{\frac{3}{m^4} + 2\left(\frac{3}{m^4} + \frac{2}{m^2}\right) + 1} \right) \\ &= \frac{1}{2} \left(1 + \frac{3}{m^2} + \sqrt{\frac{9}{m^4} + \frac{4}{m^2} + 1} \right). \end{aligned} \quad \square$$

4. NUMERICAL ESTIMATION

As explained in the previous section, to prove that the three-dimensional Gauss algorithm is exponentially strongly convergent almost everywhere it is enough to show that

$$I_n := \frac{1}{n} \int_{\Delta^3} \log \|D_n(\omega)\| \mu(d\omega) < 0$$

for some $n \in \mathbb{N}$. In this section we will discuss how an upper bound for the integral I_n can be found.

1. Choice of norm. The integral I_n obviously depends on which norm on the space of linear maps $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ is used. Any norm which is defined by

$$\|D\| = \sup_{\mathbf{v} \in \mathbb{R}^3 \setminus \{0\}} \frac{\|D\mathbf{v}\|}{\|\mathbf{v}\|} \tag{4-1}$$

for some norm $\|\cdot\|$ on \mathbb{R}^3 can be used. In our calculations we will use the norm defined by (4-1) where $\|\cdot\|$ is the standard Euclidean norm on \mathbb{R}^3 , i.e. $\|\mathbf{v}\| = \|(v_1, v_2, v_3)\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$. Recall that the norm of a linear map $D : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ induced by the Euclidean norm on \mathbb{R}^3 is given by

$$\|D\|^2 = \max\{\gamma : \gamma \text{ is an eigenvalue of } D^t D\}.$$

This norm has the disadvantage that it is hard to compute, in comparison to many other norms. Nevertheless, this norm will be used in our calculations because Monte-Carlo integration shows that the value of n required to obtain

$$\int_{\Delta^3} \log \|D_n(\boldsymbol{\omega})\| \mu(d\boldsymbol{\omega}) < 0$$

using the Euclidean norm is smaller than the value needed using any of the other standard norms, such as $\|\mathbf{v}\| = \sum_{i=1}^3 |v_i|$ or $\|\mathbf{v}\| = \max_{1 \leq i \leq 3} |v_i|$. The advantage of using a smaller value of n outweighs the disadvantage that the norm is harder to compute.

2. Choice of n . Firstly, it is necessary to find n such that $I_n < 0$. This can easily be done by estimating I_n non-rigorously using a Monte-Carlo type technique. There are two conflicting factors which determine the choice of n . On one hand, it is obviously desirable to have n as small as possible. On the other hand, the larger I_n is in absolute value, the larger our error terms in the estimation procedure can be. In practice, a compromise has to be reached between these two factors. Table 1 contains Monte-Carlo estimates for the integrals I_n .

n	I_n
1	0.16
2	0.068
3	0.035
4	0.012
5	-0.004
6	-0.015
7	-0.024
8	-0.032
9	-0.037

TABLE 1. The approximate value of the integral I_n for $n = 1, 2, \dots, 9$.

Our calculation will show that $I_8 < 0$ since $n = 8$ seems to be the best compromise.

3. Constant factors. Recall that the invariant measure μ is given by

$$\mu(d\boldsymbol{\omega}) = \frac{\rho(\boldsymbol{\omega})}{K} d\boldsymbol{\omega}.$$

Thus

$$I_8 = \frac{1}{8K} \int_{\Delta^3} \log \|D_8(\boldsymbol{\omega})\| \rho(\boldsymbol{\omega}) d\boldsymbol{\omega}.$$

Let $\gamma(D)$ denote the largest eigenvalue of $D^t D$. Then

$$\begin{aligned} I_8 &= \frac{1}{8K} \int_{\Delta^3} \log \sqrt{\gamma(D_8(\boldsymbol{\omega}))} \rho(\boldsymbol{\omega}) d\boldsymbol{\omega} \\ &= \frac{1}{16K} \int_{\Delta^3} \log (\gamma(D_8(\boldsymbol{\omega}))) \rho(\boldsymbol{\omega}) d\boldsymbol{\omega}. \end{aligned}$$

We now describe how the integral

$$\widehat{I}_8 = \int_{\Delta^3} \log (\gamma(D_8(\boldsymbol{\omega}))) \rho(\boldsymbol{\omega}) d\boldsymbol{\omega}$$

can be estimated.

4. Partitioning of Δ^3 . We will estimate the integral \widehat{I}_8 by considering Δ^3 as the union of the simplices $\Delta_{(m_1, j_1), \dots, (m_8, j_8)}$ and estimating the integral over each of these simplices separately. Of course, there are infinitely many simplices $\Delta_{(m_1, j_1), \dots, (m_8, j_8)}$, so the set of these simplices is split into two parts: a finite part, where we consider each simplex individually, and an infinite part where we use a crude upper bound for the integral.

Let Ξ_8 denote the set of elements of the 8th level of the Markov partition, i.e.,

$$\begin{aligned} \Xi_8 &= \{\Delta_{(m_1, j_1), \dots, (m_8, j_8)} : m_i \in \mathbb{N} \text{ and} \\ &\quad 1 \leq j_i \leq 3 \text{ for } 1 \leq i \leq 8\}. \end{aligned}$$

The finite part of Ξ_8 which is used for integration is denoted Z .

5. The choice of Z . The set

$$\Omega = \bigcup_{\Delta \in Z} Z$$

should have as large a measure as possible. Consequently, the elements of Z should be chosen to be as large in measure as possible. There are two general principles which can be used in the choice of Z . Firstly, if all the m_i are relatively small then the measure of $\Delta_{(m_1, j_1), \dots, (m_8, j_8)}$ will be relatively large. Secondly, if two or more m_i are large then the measure will be small.

6. Splitting of the elements of Z . The set Z is divided into three parts, $Z = Z^{(1)} \cup Z^{(2)} \cup Z^{(3)}$, according to the relative measures of the simplices. The simplices in $Z^{(1)}$ are of relatively small measure; those in $Z^{(3)}$ are of relatively large measure while $Z^{(2)}$ contains simplices of intermediate measure. Since the integral over a simplex will be estimated in terms of the values of the integrand at the vertices, it follows that, in general, the larger the simplex the less accurate the estimate will be. Consequently the simplices in $Z^{(2)}$ and $Z^{(3)}$ will be split into subsimplices and the integral over each of these subsimplices will be estimated individually. Note that the vertices of a simplex $\Delta \in Z$ can be calculated using Proposition 3.5.

A simplex $\Delta \in Z^{(2)}$ with vertices $\mathbf{v}_1, \dots, \mathbf{v}_4$ is divided into four pieces in the following way. Define

$$\mathbf{c} = \frac{1}{4}(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4).$$

Then

$$\Delta = \bigcup_{i=1}^4 \Delta_i$$

where Δ_1 is the simplex with vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{c}$; Δ_2 has vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{c}$; Δ_3 has vertices $\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4, \mathbf{c}$; and Δ_4 has vertices $\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{c}$. Let $S_4(\Delta) = \{\Delta_1, \Delta_2, \Delta_3, \Delta_4\}$.

The simplices in $Z^{(3)}$ are divided into 16 pieces. Firstly, $\Delta \in Z^{(3)}$ is split into four pieces in the same way as described above. Each of these four pieces is split into four pieces in the same way. Thus, if $\Delta \in Z^{(3)}$ let $S_4(\Delta) = \{\Delta_1, \Delta_2, \Delta_3, \Delta_4\}$. Then, if $S_{16}(\Delta)$ denotes the 16 subsimplices that Δ is split into, we have

$$S_{16}(\Delta) = \bigcup_{i=1}^4 S_4(\Delta_i).$$

Let

$$\bar{Z} = Z^{(1)} \cup \left(\bigcup_{\Delta \in Z^{(2)}} S_4(\Delta) \right) \cup \left(\bigcup_{\Delta \in Z^{(3)}} S_{16}(\Delta) \right).$$

Then

$$\Omega = \bigcup_{\Delta \in \bar{Z}} \Delta.$$

7. Estimation of the integral over $\Delta \in \bar{Z}$ and the measure of Δ . Let $\mathbf{v}_1, \dots, \mathbf{v}_4$ be the vertices of $\Delta \in \bar{Z}$.

(i) For $i = 1, \dots, 4$, calculate the matrix $D_8(\mathbf{v}_i)$ using Proposition 3.4.

(ii) Calculate $d_n(\Delta) = \log \left(\max_{1 \leq i \leq 4} \gamma(D_8(\mathbf{v}_i)) \right)$.

(iii) Find $\rho(\mathbf{v}_i)$ using (2–5) and let $\bar{\rho}(\Delta) = \max_{1 \leq i \leq 4} \rho(\mathbf{v}_i)$.

(iv) Calculate a lower bound for the density over Δ using Corollary 3.10.

(v) Find the volume of Δ using

$$\text{vol}(\Delta) = \frac{1}{6} \left| \det \begin{pmatrix} \mathbf{v}_2 - \mathbf{v}_1 \\ \mathbf{v}_3 - \mathbf{v}_1 \\ \mathbf{v}_4 - \mathbf{v}_1 \end{pmatrix} \right|.$$

(vi) Let

$$\bar{T}^{(u)}(\Delta) = \begin{cases} d_n(\Delta) \underline{\rho}(\Delta) \text{vol}(\Delta) & \text{if } d_n(\Delta) < 0; \\ d_n(\Delta) \bar{\rho}(\Delta) \text{vol}(\Delta) & \text{if } d_n(\Delta) > 0. \end{cases}$$

Notice that

$$\int_{\Delta} \log(\gamma(D_8(\boldsymbol{\omega}))) \rho(\boldsymbol{\omega}) d\boldsymbol{\omega} \leq \bar{T}^{(u)}(\Delta).$$

The symbol “(u)” is intended to denote that this is an upper bound for the “unnormalised” integral. Note that

$$\frac{1}{8} \int_{\Delta} \log \|D_8(\boldsymbol{\omega})\| \mu(d\boldsymbol{\omega}) \leq \frac{1}{16K} \bar{T}^{(u)}(\Delta).$$

(vii) Let $\underline{\mu}^{(u)}(\Delta) = \underline{\rho}(\Delta) \text{vol}(\Delta)$. Then

$$\int_{\Delta} \rho(\boldsymbol{\omega}) d\boldsymbol{\omega} \geq \underline{\mu}^{(u)}(\Delta).$$

8. Estimation of the integral over Ω and the measure of Ω . The following estimates hold:

$$\int_{\Omega} \log(\gamma(D_8(\boldsymbol{\omega}))) \rho(\boldsymbol{\omega}) d\boldsymbol{\omega} \leq \sum_{\Delta \in \bar{Z}} \bar{T}^{(u)}(\Delta)$$

and

$$\int_{\Omega} \rho(\boldsymbol{\omega}) d\boldsymbol{\omega} \geq \sum_{\Delta \in \bar{Z}} \underline{\mu}^{(u)}(\Delta).$$

9. An upper bound for the integral over the complementary set $\Delta^3 \setminus \Omega$. By Lemma 3.12, $\|D(T\boldsymbol{\omega})D(\boldsymbol{\omega})\| \leq \sqrt{7}$. Thus $\|D_8(\boldsymbol{\omega})\| \leq (\sqrt{7})^4 = 49$, which implies that

$$\begin{aligned} \int_{\Delta^3 \setminus \Omega} \log \|D_8(\boldsymbol{\omega})\| \mu(d\boldsymbol{\omega}) &\leq \log(49) \mu(\Delta^3 \setminus \Omega) \\ &= \log(49) (1 - \mu(\Omega)) \\ &\leq \log(49) \left(1 - \frac{1}{K} \sum_{\Delta \in \bar{Z}} \underline{\mu}^{(u)}(\Delta) \right). \end{aligned}$$

Set of simplices	All simplices $\Delta_{(m_1, j_1), \dots, (m_8, j_8)}$ for which $1 \leq j_i \leq 3$ for $i = 1, \dots, 8$ and:
Z_1	$\#\{i : 1 \leq m_i \leq 3\} = 8$ or $\#\{i : 1 \leq m_i \leq 3\} = 7$ and $\#\{i : 4 \leq m_i \leq 6\} = 1$
Z_2	$\#\{i : 7 \leq m_i \leq 14\} = 1, \quad \#\{i : 1 \leq m_i \leq 3\} = 7$
Z_3	$\#\{i : 15 \leq m_i \leq 22\} = 1, \quad \#\{i : 1 \leq m_i \leq 3\} = 7$
Z_4	$\#\{i : 4 \leq m_i \leq 6\} = 2, \quad \#\{i : 1 \leq m_i \leq 3\} = 6$
Z_5	$\#\{i : 4 \leq m_i \leq 6\} = 3, \quad \#\{i : 1 \leq m_i \leq 3\} = 5$
Z_6	$\#\{i : 4 \leq m_i \leq 6\} = 4, \quad \#\{i : 1 \leq m_i \leq 3\} = 4$
Z_7	$\#\{i : 4 \leq m_i \leq 6\} = 5, \quad \#\{i : 1 \leq m_i \leq 3\} = 3$
Z_8	$\#\{i : 4 \leq m_i \leq 6\} = 6, \quad \#\{i : 1 \leq m_i \leq 3\} = 2$
Z_9	$\#\{i : 4 \leq m_i \leq 6\} = 7, \quad \#\{i : 1 \leq m_i \leq 3\} = 1$
Z_{10}	$\#\{i : 4 \leq m_i \leq 6\} = 8$
Z_{11}	$\#\{i : 4 \leq m_i \leq 6\} = 1, \quad \#\{i : 7 \leq m_i \leq 9\} = 1, \quad \#\{i : 1 \leq m_i \leq 3\} = 6$

TABLE 2. The definition of Z_1, \dots, Z_{11} .

10. An upper bound for I_8 . We have the following final upper bound for I_8 :

$$\begin{aligned}
 I_8 &= \frac{1}{8} \int_{\Delta^3} \log \|D_8(\omega)\| \mu(d\omega) \\
 &= \frac{1}{16K} \int_{\Omega} \log(\gamma(D_8(\omega))) \rho(\omega) d\omega \\
 &\quad + \frac{1}{8} \int_{\Delta^3 \setminus \Omega} \log \|D_8(\omega)\| \mu(d\omega) \\
 &\leq \frac{1}{16K} \left(\sum_{\Delta \in \bar{Z}} \bar{T}^{(u)}(\Delta) \right) + \frac{1}{4} \log(7) \left(1 - \frac{1}{K} \sum_{\Delta \in \bar{Z}} \underline{\mu}^{(u)}(\Delta) \right).
 \end{aligned}$$

5. NUMERICAL RESULTS

The method described in the previous section can easily be implemented on a computer. In this section we will present numerical results which were obtained using this method and which prove that

$$I_8 = \frac{1}{8} \int_{\Delta^3} \log \|D_8(\omega)\| \mu(d\omega) < 0.$$

For convenience of presentation of these results, we split the set Z of simplices which are used for integration into 11 parts, Z_1, Z_2, \dots, Z_{11} . These sets are defined in Table 2.

The simplices in Z_1 were split into 16 pieces, those in Z_2 were split into 4 pieces, and the remaining simplices were not split at all, i.e.

$$Z^{(3)} = Z_1, \quad Z^{(2)} = Z_2, \quad Z^{(1)} = \bigcup_{i=3}^{12} Z_i.$$

Let

$$\Omega_i = \bigcup_{\Delta \in Z_i} \Delta.$$

Table 3 contains the numerical results. The program which produced these results was written in C. To evaluate $\gamma(D_8(\omega))$, which requires the calculation of the largest eigenvalue of the symmetric matrix $(D_8(\omega))^t D_8(\omega)$, the routines tred2 and tqli (see pages 474 and 480 of [Press et al. 92]) were used.

i	Upper bound for the integral over Ω_i , i.e. $\frac{1}{16K} \sum_{\Delta \in Z_i} \bar{T}^{(u)}(\Delta)$	Lower bound for the measure of Ω_i , i.e. $\frac{1}{K} \sum_{\Delta \in Z_i} \underline{\mu}^{(u)}(\Delta)$
1	-0.0155527	0.866392
2	-6.58355×10^{-4}	0.0610718
3	-1.15405×10^{-5}	6.09373×10^{-3}
4	-5.94629×10^{-4}	0.0281124
5	-6.53896×10^{-5}	1.65073×10^{-3}
6	-3.24197×10^{-6}	4.72975×10^{-5}
7	-7.14697×10^{-8}	6.80746×10^{-7}
8	-7.30254×10^{-10}	4.95039×10^{-9}
9	-3.37587×10^{-12}	1.72503×10^{-11}
10	-5.60129×10^{-15}	2.28312×10^{-14}
11	-1.85439×10^{-4}	0.0106883
Total	-0.0170713	0.974056

TABLE 3. Results for regions $\Omega_1, \dots, \Omega_{11}$.

The first of these routines transforms a real, symmetric matrix into tridiagonal form by a series of orthogonal transformations. The second routine finds the eigenvalues of such a matrix. (Quicker algorithms to do this may exist; a referee suggested that it might be better for future work to use the singular value decomposition to

compute the norm rather than computing the eigenvalues, and noted that there is a variety of public domain software for both eigenvalues and singular values, notably LAPACK.)

The evaluation of the matrix $V_n = \tilde{A}_{(m_1, j_1)} \tilde{A}_{(m_2, j_2)} \cdots \tilde{A}_{(m_n, j_n)} V_0$ (see Proposition 3.5) was performed using integer variables. All other calculations were performed to double precision. The program took approximately 750 hours on a Sun Ultra 5 to produce the given results. The program was also written in the Matlab language using the built-in eigenvalue and matrix multiplication routines. Many of the calculations were repeated using this program and the results agreed with the ones obtained by the C program to the accuracy stated.

Proposition 5.1. *For any $\omega \in \Delta^3 \setminus \Omega$*

$$\frac{1}{8} \log \|D_8(\omega)\| \leq 0.452590. \tag{5-1}$$

Proof: Since the set Ω contains all simplices $\Delta_{(m_1, j_1), \dots, (m_8, j_8)}$ for which $1 \leq m_i \leq 6$ for $1 \leq i \leq 8$, if $\omega \in \Delta^3 \setminus \Omega$ then for some $1 \leq l \leq 8$, $m_l \geq 7$. Consequently, by Corollary 3.14,

$$\|D(T^{l-1}\omega)\| \leq \left(\frac{1}{2} \left(1 + \frac{3}{7^2} + \sqrt{\frac{9}{7^4} + \frac{4}{7^2} + 1}\right)\right)^{\frac{1}{2}}.$$

Also, since $m(\omega) \geq 1$ for any ω , for $1 \leq i \leq 8$

$$\|D(T^{i-1}\omega)\|^2 \leq \frac{1}{2} \left(4 + \sqrt{14}\right). \tag{5-2}$$

The matrix product $D(T^7\omega) \cdots D(T\omega)D(\omega)$ can be viewed as the product of three matrices of the form $D(T^i\omega)D(T^{i-1}\omega)$, one matrix $D(T^{i-1}\omega)$ for which (5-2) holds, and the matrix $D(T^{l-1}\omega)$. Thus, by Lemma 3.12,

$$\begin{aligned} \|D_8(\omega)\| &\leq (\sqrt{7})^3 \left(\frac{1}{2} \left(1 + \frac{3}{7^2} + \sqrt{\frac{9}{7^4} + \frac{4}{7^2} + 1}\right)\right)^{\frac{1}{2}} \\ &\quad \times \left(\frac{1}{2} \left(4 + \sqrt{14}\right)\right)^{\frac{1}{2}} \\ &\leq 37.364389 \end{aligned}$$

This implies (5-1). □

Theorem 5.2.

$$\frac{1}{8} \int_{\Delta^3} \log \|D_8(\omega)\| \mu(d\omega) \leq -0.00532930$$

Proof: By Table 3 and Proposition 5.1

$$\begin{aligned} \frac{1}{8} \int_{\Delta^3} \log \|D_8(\omega)\| \mu(d\omega) &\leq -0.0170713 \\ &\quad + 0.452590(1 - 0.974056) \\ &= -0.00532930. \end{aligned} \quad \square$$

Corollary 5.3. *The three-dimensional Gauss algorithm is exponentially strongly convergent almost everywhere.*

6. DISCUSSION

Similar calculations to those described above can be carried out in any dimension. However, the number of calculations necessary increases very rapidly with the dimension. One reason for this is that the value of n necessary to have $I_n < 0$ increases with the dimension. In dimension 3, $I_n < 0$ for $n \geq 5$, whereas in four dimensions, $I_n < 0$ for $n \geq 12$. This obviously increases the number of calculations necessary. Another problem is that $\lambda_1(D) \rightarrow 0$ as the dimension increases. This means that the error terms must be smaller, so more calculations will be necessary. The calculation can, of course, be broken down into several disjoint calculations which can then be carried out simultaneously on different computers. More modern machines than the one used in this paper may be of the order of 10 times quicker. Using several of these machines would result in a 100 times improvement in the speed of the calculation. However, even allowing for such an improvement, I believe that the calculations necessary to prove strong convergence in dimension 4 would still take a great deal longer than the calculations described here. There seems to be no reason to doubt that the result is true in all dimensions.

It should be possible to prove strong convergence of other three-dimensional continued fraction algorithms by similar methods to those in this paper, if the invariant density is known explicitly. If a good approximation to the invariant density is known then it should also be possible to prove the results. However, obtaining such an approximation seems to be very difficult.

ACKNOWLEDGMENTS

This paper is the natural continuation of a series of papers by myself and K. Khanin [Hardcastle, Khanin 00], [Hardcastle, Khanin 01], [Hardcastle, Khanin 02]. I am very grateful to K. Khanin for numerous useful discussions concerning the work in this paper. I am also grateful to R. Khanin for advice about computational matters. I also thank the Engineering and Physical Sciences Research Council of the UK for financial support.

REFERENCES

- [Baldwin 92a] P. R. Baldwin. “A multidimensional continued fraction and some of its statistical properties,” *Jour. Stat. Phys.* **66** (1992), 1463–1505.
- [Baldwin 92b] P. R. Baldwin. “A convergence exponent for multidimensional continued-fraction algorithms,” *Jour. Stat. Phys.* **66** (1992), 1507–1526.
- [Brun 57] V. Brun. “Algorithmes euclidiens pour trois et quatre nombres,” In *13 ième Congre. Math. Scand., Helsinki* (1957), 45–64.
- [Ito et al. 96] T. Fujita, S. Ito, M. Keane and M. Ohtsuki. “On almost everywhere exponential convergence of the modified Jacobi-Perron algorithm: a corrected proof,” *Ergod. Th. and Dyn. Sys.* **16** (1996), 1345–1352.
- [Hardcastle, Khanin 00] D. M. Hardcastle and K. Khanin. “On almost everywhere strong convergence of multidimensional continued fraction algorithms,” *Ergod. Th. and Dyn. Sys.* **20** (2000), 1711–1733.
- [Hardcastle, Khanin 01] D. M. Hardcastle and K. Khanin. “Continued fractions and the d -dimensional Gauss transformation,” *Commun. Math. Phys.* **215** (2001), 487–515.
- [Hardcastle, Khanin 02] D. M. Hardcastle and K. Khanin. “The d -dimensional Gauss transformation: strong convergence and Lyapunov exponents,” *Experimental Mathematics* **11:1** (2002), 119–129.
- [Ito et al. 93] S. Ito, M. Keane and M. Ohtsuki. “Almost everywhere exponential convergence of the modified Jacobi-Perron algorithm,” *Ergod. Th. and Dyn. Sys.* **13** (1993), 319–334.
- [Jacobi 1868] C. G. J. Jacobi. “Allgemeine Theorie der kettenbruchähnlichen Algorithmen, in welchen jede Zahl aus drei vorhergehenden gebildet wird,” *J. Reine Angew. Math.* **69** (1868), 29–64.
- [Khanin 92] K. Khanin. Talk at the International Workshop on Dynamical Systems, Porto (1992).
- [Kosygin 91] D. V. Kosygin. “Multidimensional KAM theory from the renormalization group viewpoint,” In *Dynamical Systems and Statistical Mechanics (Ed.: Ya. G. Sinai), Advances in Soviet Mathematics* **3** (1991), 99–129.
- [Kingman 68] J. F. C. Kingman. “The ergodic theory of sub-additive stochastic processes,” *J. Royal Stat. Soc. B* **30** (1968), 499–510.
- [Lagarias 93] J. C. Lagarias. “The quality of the Diophantine approximations found by the Jacobi-Perron algorithm and related algorithms,” *Mh. Math.* **115** (1993), 299–328.
- [Meester 98] R. Meester. “A simple proof of the exponential convergence of the modified Jacobi-Perron algorithm,” *Ergod. Th. and Dyn. Sys.* **19** (1999), 1077–1083.
- [Oseledets 68] V. I. Oseledets. “A multiplicative ergodic theorem. Liapunov characteristic numbers for dynamical systems,” *Trans. Moscow Math. Soc.* **19** (1968), 197–221.
- [Paley, Ursell 30] R. E. A. C. Paley and H. D. Ursell. “Continued fractions in several dimensions,” *Proc. Cambridge Philos. Soc.* **26** (1930), 127–144.
- [Perron 1907] O. Perron. “Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus,” *Math. Ann.* **64** (1907), 1–76.
- [Podsypanin 77] E. V. Podsypanin. “A generalization of the algorithm for continued fractions related to the algorithm of Viggo Brun,” *Zap. Nauchn. Sem. Leningrad Otdel. Mat. Inst. Steklov* **67** (1977), 184–194. English translation: *Journal of Soviet Math.* **16** (1981), 885–893.
- [Press et al. 92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing, Second Edition*, Cambridge University Press (1992).
- [Schratzberger 98] B. R. Schratzberger. “The exponent of convergence for Brun’s algorithm in two dimensions,” *Sitzungsber. Abt. II* **207** (1998), 229–238.
- [Schweiger 79] F. Schweiger. “A modified Jacobi-Perron algorithm with explicitly given invariant measure,” In *Ergodic Theory, Proceedings Oberwolfach, Germany 1978, Lecture Notes in Mathematics* **729** (1979), 199–202, Springer-Verlag.
- [Schweiger 96] F. Schweiger. “The exponent of convergence for the 2-dimensional Jacobi-Perron algorithm,” In *Proceedings of the Conference on Analytic and Elementary Number Theory in Vienna 1996 (Ed.: W. G. Nowak and J. Schoissengeier)*, 207–213.
- [Selmer 61] E. Selmer. “Om flerdimensjonal Kjede brøk,” *Nordisk Mat. Tidskr.* **9** (1961), 37–43.

D. M. Hardcastle, Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS,
United Kingdom (D.M.Hardcastle@ma.hw.ac.uk)

Received February 28, 2001; accepted in revised form August 14, 2001.